

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 May 2001 (17.05.2001)

PCT

(10) International Publication Number
WO 01/34804 A1

(51) International Patent Classification⁷: C12N 15/12,
15/11, 15/62, C07K 14/705, 16/28, A61K 38/17, G01N
33/68

(74) Agents: ISHIMOTO, Lance et al.; Lexicon Genetics In-
corporated, 4000 Research Forest Drive, The Woodlands,
TX 77381 (US).

(21) International Application Number: PCT/US00/31045

(22) International Filing Date:
10 November 2000 (10.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/165,263 12 November 1999 (12.11.1999) US

(71) Applicant: LEXICON GENETICS INCORPORATED
[US/US]; 4000 Research Forest Drive, The Woodlands, TX
77381 (US).

(72) Inventors: TURNER, C., Alexander, Jr.; 67 Win-
ter Wheat Place, The Woodlands, TX 77381 (US).
MATHUR, Brian; 12000 Sawmill Road #2014, The
Woodlands, TX 77380 (US). FRIEDRICH, Glenn; Bre-
land & Breland, 2207 Hermann Drive, Houston, TX 77004
(US). ZAMBROWICZ, Brian; 18 Firethorne Place, The
Woodlands, TX 77382 (US). SANDS, Arthur, T.; 163
Bristol Bend Circle, The Woodlands, TX 77382 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- With international search report.
- Before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments.

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: HUMAN MEMBRANE PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

(57) Abstract: Human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and
pharmacogenomic applications.

WO 01/34804 A1

HUMAN MEMBRANE PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

The present application claims the benefit of U.S. Provisional Application Number 60/165,263 which was filed on November 12, 1999 and is herein incorporated by reference in its entirety.

1. INTRODUCTION

The present invention relates to the discovery, identification, and characterization of novel human polynucleotides encoding proteins that share sequence similarity with animal MHC proteins. The invention encompasses the described polynucleotides, host cell expression systems, the encoded proteins, fusion proteins, polypeptides and peptides, antibodies to the encoded proteins and peptides, and genetically engineered animals that either lack or over express the disclosed genes, antagonists and agonists of the proteins, and other compounds that modulate the expression or activity of the proteins encoded by the disclosed genes that can be used for diagnosis, drug screening, clinical trial monitoring, the treatment of diseases and disorders, or otherwise contributing to the quality of life.

2. BACKGROUND OF THE INVENTION

Membrane proteins play important roles as, *inter alia*, cell surface markers, receptors, and mediators of signal transduction.

3. SUMMARY OF THE INVENTION

The present invention relates to the discovery, identification, and characterization of nucleotides that encode novel human proteins, and the corresponding amino acid sequences of these proteins. The novel human proteins (NHPs) described for the first time herein share structural similarity with membrane receptors such as, but not limited to, MHC or HLA proteins.

The novel human nucleic acid sequences described herein, encode alternative proteins/open reading frames (ORFs) of 585 and 584 amino acids in length (see SEQ ID NOS: 2 and 4 respectively).

5 The invention also encompasses agonists and antagonists of the described NHPs, including small molecules, large molecules, mutant NHPs, or portions thereof that compete with native NHP, peptides, and antibodies, as well as nucleotide sequences that can be used to inhibit the expression of the
10 described NHPs (e.g., antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs) or to enhance the expression of the described NHP sequences (e.g., expression constructs that place the described sequence under the control of a strong promoter system), and transgenic
15 animals that express a NHP transgene, or "knock-outs" (which can be conditional) that do not express a functional NHP.

Further, the present invention also relates to processes for identifying compounds that modulate, i.e., act as agonists or antagonists, of NHP expression and/or NHP activity that
20 utilize purified preparations of the described NHPs and/or NHP product, or cells expressing the same. Such compounds can be used as therapeutic agents for the treatment of any of a wide variety of symptoms associated with biological disorders or imbalances.

25

4. DESCRIPTION OF THE SEQUENCE LISTING AND FIGURES

The Sequence Listing provides the sequences of the several NHP ORFs that encode the described NHP amino acid sequences.

30

5. DETAILED DESCRIPTION OF THE INVENTION

The NHPs, described for the first time herein, are novel proteins that are expressed in, *inter alia*, human cell lines, and human kidney, bone marrow, adrenal, trachea, brain,
35 pancreas, mammary gland, placenta, prostate, thymus, liver, lymph node, testis, and gene trapped human cells. Gene

trapped knockout ES cell lines have been produced that bear mutations in the murine ortholog encoding the described NHPs.

The present invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, the expression products of such nucleotides, and:

5 (a) nucleotides that encode mammalian homologs of the described sequences, including the specifically described NHPs, and the NHP products; (b) nucleotides that encode one or more portions of the NHPs that correspond to functional

10 domains, and the polypeptide products specified by such nucleotide sequences, including but not limited to the novel regions of any active domain(s); (c) isolated nucleotides that encode mutant versions, engineered or naturally occurring, of the described NHPs in which all or a part of at least one

15 domain is deleted or altered, and the polypeptide products specified by such nucleotide sequences, including but not limited to soluble proteins and peptides in which all or a portion of the signal sequence is deleted; (d) nucleotides that encode chimeric fusion proteins containing all or a

20 portion of a coding region of an NHP, or one of its domains (e.g., a receptor or ligand binding domain, accessory protein/self-association domain, etc.) fused to another peptide or polypeptide; or (e) therapeutic or diagnostic derivatives of the described polynucleotides such as

25 oligonucleotides, antisense polynucleotides, ribozymes, dsRNA, or gene therapy constructs comprising a sequence first disclosed in the Sequence Listing.

As discussed above, the present invention includes:

(a) the human DNA sequences presented in the Sequence Listing

30 (and vectors comprising the same) and additionally contemplates any nucleotide sequence encoding a contiguous NHP open reading frame (ORF) that hybridizes to a complement of a DNA sequence presented in the Sequence Listing under highly stringent conditions, e.g., hybridization to filter-bound DNA

35 in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol.

I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3) and encodes a functionally equivalent gene product. Additionally contemplated are any nucleotide sequences that hybridize to the complement of a DNA
5 sequence that encodes and expresses an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42 °C (Ausubel et al., 1989, *supra*), yet still encodes a functionally equivalent NHP product. Functional equivalents of a NHP
10 include naturally occurring NHPs present in other species and mutant NHPs whether naturally occurring or engineered (by site directed mutagenesis, gene shuffling, directed evolution as described in, for example, U.S. Patent No. 5,837,458). The invention also includes degenerate nucleic acid variants of
15 the disclosed NHP polynucleotide sequences.

Additionally contemplated are polynucleotides encoding NHP ORFs, or their functional equivalents, encoded by polynucleotide sequences that are about 99, 95, 90, or about 85 percent similar or identical to corresponding regions of
20 the nucleotide sequences of the Sequence Listing (as measured by BLAST sequence comparison analysis using, for example, the GCG sequence analysis package (Madison, Wisconsin) using standard default settings).

The invention also includes nucleic acid molecules,
25 preferably DNA molecules, that hybridize to, and are therefore the complements of, the described NHP nucleotide sequences. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances where the nucleic acid molecules are deoxyoligonucleotides ("DNA
30 oligos"), such molecules are generally about 16 to about 100 bases long, or about 20 to about 80, or about 34 to about 45 bases long, or any variation or combination of sizes represented therein that incorporate a contiguous region of sequence first disclosed in the Sequence Listing. Such
35 oligonucleotides can be used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and sequencing templates, etc.

Alternatively, such NHP oligonucleotides can be used as hybridization probes for screening libraries, and assessing gene expression patterns (particularly using a micro array or high-throughput "chip" format). Additionally, a series of the described NHP oligonucleotide sequences, or the complements thereof, can be used to represent all or a portion of the described NHP sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length can partially overlap each other and/or the NHP sequence may be represented using oligonucleotides that do not overlap. Accordingly, the described NHP polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide sequences of at least about 18, and preferably about 25, nucleotides in length that are each first disclosed in the described Sequence Listing. Such oligonucleotide sequences may begin at any nucleotide present within a sequence in the Sequence Listing and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

For oligonucleotide probes, highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as NHP gene antisense molecules, useful, for example, in NHP gene regulation (for and/or as antisense primers in amplification reactions of NHP gene nucleic acid sequences). With respect to NHP gene regulation, such techniques can be used to regulate biological functions. Further, such sequences may be used as part of ribozyme and/or triple helix sequences that are also useful for NHP gene regulation.

Inhibitory antisense or double stranded oligonucleotides can additionally comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine,

5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-
 2-thiouridine, 5-carboxymethylaminomethyluracil,
 dihydrouracil, beta-D-galactosylqueosine, inosine,
 N6-isopentenyladenine, 1-methylguanine, 1-methylinosine,
 5 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine,
 3-methylcytosine, 5-methylcytosine, N6-adenine,
 7-methylguanine, 5-methylaminomethyluracil,
 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine,
 5'-methoxycarboxymethyluracil, 5-methoxyuracil,
 10 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid
 (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine,
 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil,
 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-
 5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-
 15 carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

20 In yet another embodiment, the antisense oligonucleotide will comprise at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl
 25 phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an α -anomeric oligonucleotide. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the
 30 strands run parallel to each other (Gautier *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue *et al.*, 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, FEBS Lett. 215:327-330). Alternatively, double stranded
 35 RNA can be used to disrupt the expression and function of a targeted NHP.

Oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples,

5 phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

10 Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al., 1989, Molecular Cloning, A
15 Laboratory Manual (and periodic updates thereof), Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Alternatively, suitably labeled NHP nucleotide probes can
20 be used to screen a human genomic library using appropriately stringent conditions or by PCR. The identification and characterization of human genomic clones is helpful for identifying polymorphisms (including, but not limited to, nucleotide repeats, microsatellite alleles, single nucleotide
25 polymorphisms, or coding single nucleotide polymorphisms), determining the genomic structure of a given locus/allele, and designing diagnostic tests. For example, sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in
30 amplification assays to detect mutations within the exons, introns, splice sites (e.g., splice acceptor and/or donor sites), etc., that can be used in diagnostics and pharmacogenomics.

Further, a NHP sequence homolog can be isolated from
35 nucleic acid from an organism of interest by performing PCR using two degenerate or "wobble" oligonucleotide primer pools designed on the basis of amino acid sequences within the NHP

products disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express an allele of a NHP gene.

The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequence of the desired NHP gene. The PCR fragment can then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment can be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library. Alternatively, the labeled fragment can be used to isolate genomic clones via the screening of a genomic library.

PCR technology can also be used to isolate full length cDNA sequences. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*, one known, or suspected, to express a NHP gene). A reverse transcription (RT) reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" using a standard terminal transferase reaction, the hybrid may be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA sequences upstream of the amplified fragment can be isolated. For a review of cloning strategies that can be used, see *e.g.*, Sambrook *et al.*, 1989, *supra*.

A cDNA encoding a mutant NHP gene can be isolated, for example, by using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying a mutant NHP allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, optionally

cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant NHP allele to that of a corresponding normal NHP allele, the mutation(s) responsible for the loss or alteration of function of the mutant NHP gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant NHP allele (e.g., a person manifesting a NHP-associated phenotype such as, for example, obesity, high blood pressure, connective tissue disorders, infertility, etc.), or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant NHP allele. A normal NHP gene, or any suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant NHP allele in such libraries. Clones containing mutant NHP sequences can then be purified and subjected to sequence analysis according to methods well known to those skilled in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant NHP allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue can be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against a normal NHP product, as described below. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) Additionally, screening can be accomplished by screening with labeled NHP fusion proteins, such as, for example, alkaline phosphatase-NHP or NHP-alkaline phosphatase fusion proteins. In cases where a NHP mutation results in an expressed gene product with altered function (e.g., as a result of a missense or a frameshift mutation), polyclonal antibodies to a NHP are likely to cross-react with a corresponding mutant NHP gene

product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known in the art.

The invention also encompasses (a) DNA vectors that
5 contain any of the foregoing NHP coding sequences and/or their complements (*i.e.*, antisense); (b) DNA expression vectors that contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences (for example, baculo virus
10 as described in U.S. Patent No. 5,869,336 herein incorporated by reference); (c) genetically engineered host cells that contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell; and (d)
15 genetically engineered host cells that express an endogenous NHP gene under the control of an exogenously introduced regulatory element (*i.e.*, gene activation). As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other
20 elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include but are not limited to the human cytomegalovirus (hCMV) immediate early gene, regulatable, viral elements (particularly retroviral LTR promoters), the early or late promoters of SV40
25 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast α -mating
30 factors.

The present invention also encompasses antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists and agonists of the NHP, as well as compounds or nucleotide constructs that inhibit expression of a NHP gene
35 (transcription factor inhibitors, antisense and ribozyme molecules, or gene or regulatory sequence replacement constructs), or promote the expression of a NHP (*e.g.*,

expression constructs in which NHP coding sequences are operatively associated with expression control elements such as promoters, promoter/enhancers, etc.).

The NHPs or NHP peptides, NHP fusion proteins, NHP nucleotide sequences, antibodies, antagonists and agonists can be useful for the detection of mutant NHPs or inappropriately expressed NHPs for the diagnosis of disease. The NHP proteins or peptides, NHP fusion proteins, NHP nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and genetically engineered cells and animals can be used for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of the symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body. The use of engineered host cells and/or animals may offer an advantage in that such systems allow not only for the identification of compounds that bind to the endogenous receptor for an NHP, but can also identify compounds that trigger NHP-mediated activities or pathways.

Finally, the NHP products can be used as therapeutics. For example, soluble derivatives such as NHP peptides/domains corresponding the NHPs, NHP fusion protein products (especially NHP-Ig fusion proteins, i.e., fusions of a NHP, or a domain of a NHP, to an IgFc), NHP antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists or agonists (including compounds that modulate or act on downstream targets in a NHP-mediated pathway) can be used to directly treat diseases or disorders. For instance, the administration of an effective amount of soluble NHP, or a NHP-IgFc fusion protein or an anti-idiotypic antibody (or its Fab) that mimics the NHP could activate or effectively antagonize the endogenous NHP receptor. Nucleotide constructs encoding such NHP products can be used to genetically engineer host cells to express such products *in vivo*; these genetically engineered cells function as "bioreactors" in the body delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion protein to the body. Nucleotide constructs

encoding functional NHPs, mutant NHPs, as well as antisense and ribozyme molecules can also be used in "gene therapy" approaches for the modulation of NHP expression. Thus, the invention also encompasses pharmaceutical formulations and methods for treating biological disorders.

Various aspects of the invention are described in greater detail in the subsections below.

5.1 THE NHP SEQUENCES

The cDNA sequences and the corresponding deduced amino acid sequences of the described NHPs are presented in the Sequence Listing. The NHP nucleotides were obtained from clustered human gene trapped sequences, ESTs and human kidney and bone marrow cDNA libraries (Edge Biosystems, Gaithersburg, MD). The described sequences share structural similarity with class III MHC proteins and also display multiple transmembrane regions as have been seen in similar proteins. The described sequences can also incorporate a polymorphism at amino acid position 370 of, for example, SEQ ID NO:2 which changes a I to a V.

5.2 NHPS AND NHP POLYPEPTIDES

NHPs, polypeptides, peptide fragments, mutated, truncated, or deleted forms of the NHPs, and/or NHP fusion proteins can be prepared for a variety of uses. These uses include, but are not limited to, the generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to a NHP, as reagents in assays for screening for compounds that can be as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or medical disorders and disease. Given the similarity information and expression data, the described NHPs can be targeted (by drugs, oligos, antibodies, etc,) in order to treat disease, or to therapeutically augment the efficacy of therapeutic agents.

The Sequence Listing discloses the amino acid sequences encoded by the described NHP genes. The NHPs typically

display initiator methionines in DNA sequence contexts consistent with a translation initiation site, and a signal sequence characteristic of membrane or secreted proteins.

The NHP amino acid sequences of the invention include the amino acid sequences presented in the Sequence Listing as well as analogues and derivatives thereof. Further, corresponding NHP homologues from other species are encompassed by the invention. In fact, any NHP protein encoded by the NHP nucleotide sequences described above are within the scope of the invention, as are any novel polynucleotide sequences encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well known, and, accordingly, each amino acid presented in the Sequence Listing, is generically representative of the well known nucleic acid "triplet" codon, or in many cases codons, that can encode the amino acid. As such, as contemplated herein, the amino acid sequences presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at page 109 of "Molecular Cell Biology", 1986, J. Darnell et al. eds., Scientific American Books, New York, NY, herein incorporated by reference) are generically representative of all the various permutations and combinations of nucleic acid sequences that can encode such amino acid sequences.

The invention also encompasses proteins that are functionally equivalent to the NHPs encoded by the presently described nucleotide sequences as judged by any of a number of criteria, including, but not limited to, the ability to bind and cleave a substrate of a NHP, or the ability to effect an identical or complementary downstream pathway, or a change in cellular metabolism (e.g., proteolytic activity, ion flux, tyrosine phosphorylation, transport, etc.). Such functionally equivalent NHP proteins include, but are not limited to, additions or substitutions of amino acid residues within the amino acid sequence encoded by the NHP nucleotide sequences described above, but which result in a silent change, thus producing a functionally equivalent gene product. Amino acid

substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

A variety of host-expression vector systems can be used to express the NHP nucleotide sequences of the invention. Where, as in the present instance, the NHP peptide or polypeptide is thought to be membrane protein, the hydrophobic regions of the protein can be excised and the resulting soluble peptide or polypeptide can be recovered from the culture media. Such expression systems also encompass engineered host cells that express a NHP, or functional equivalent, *in situ*. Purification or enrichment of a NHP from such expression systems can be accomplished using appropriate detergents and lipid micelles and methods well known to those skilled in the art. However, such engineered host cells themselves may be used in situations where it is important not only to retain the structural and functional characteristics of the NHP, but to assess biological activity, e.g., in drug screening assays.

The expression systems that can be used for purposes of the invention include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing NHP nucleotide sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing NHP nucleotide sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing NHP sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic

virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing NHP nucleotide sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs

5 containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may
10 be advantageously selected depending upon the use intended for the NHP product being expressed. For example, when a large quantity of such a protein is to be produced for the generation of pharmaceutical compositions of or containing NHP, or for raising antibodies to a NHP, vectors that direct
15 the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which a NHP coding sequence may be ligated individually into the vector in frame
20 with the *lacZ* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Pharmacia or American Type Culture Collection) can also be used to express
25 foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The PGEX vectors are
30 designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhidrosis virus (AcNPV) is used as a vector to express
35 foreign genes. The virus grows in *Spodoptera frugiperda* cells. A NHP gene coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene)

of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of NHP gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (e.g., see Smith et al., 1983, J. Virol. 46:584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the NHP nucleotide sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a NHP product in infected hosts (e.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted NHP nucleotide sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire NHP gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of a NHP coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription

terminators, etc. (See Bittner *et al.*, 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or
5 modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the
10 post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper
15 processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include, but are not limited to, CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, human cell lines.

20 For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the NHP sequences described above can be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be
25 transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in
30 an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.
35 This method may advantageously be used to engineer cell lines which express the NHP product. Such engineered cell lines may be particularly useful in screening and evaluation of

compounds that affect the endogenous activity of the NHP product.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk⁻, hgprt⁻ or aprt⁻ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hygromycin (Santerre, et al., 1984, Gene 30:147).

Alternatively, any fusion protein can be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni²⁺-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

5.3 ANTIBODIES TO NHP PRODUCTS

Antibodies that specifically recognize one or more epitopes of a NHP, or epitopes of conserved variants of a NHP, or peptide fragments of a NHP are also encompassed by the

invention. Such antibodies include but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention may be used, for example, in the detection of NHP in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of NHP. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes for the evaluation of the effect of test compounds on expression and/or activity of a NHP gene product. Additionally, such antibodies can be used in conjunction gene therapy to, for example, evaluate the normal and/or engineered NHP-expressing cells prior to their introduction into the patient. Such antibodies may additionally be used as a method for the inhibition of abnormal NHP activity. Thus, such antibodies may, therefore, be utilized as part of treatment methods.

For the production of antibodies, various host animals may be immunized by injection with the NHP, an NHP peptide (e.g., one corresponding to a functional domain of an NHP), truncated NHP polypeptides (NHP in which one or more domains have been deleted), functional equivalents of the NHP or mutated variant of the NHP. Such host animals may include but are not limited to pigs, rabbits, mice, goats, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's adjuvant (complete and incomplete), mineral salts such as aluminum hydroxide or aluminum phosphate, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Alternatively, the immune response could be enhanced by combination and or coupling with molecules such as keyhole

limpet hemocyanin, tetanus toxoid, diphtheria toxoid, ovalbumin, cholera toxin or fragments thereof. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

5 Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and
10 Milstein, (1975, *Nature* 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, *Immunology Today* 4:72; Cole et al., 1983, *Proc. Natl. Acad. Sci. USA* 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, *Monoclonal Antibodies And Cancer*
15 *Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the
20 presently preferred method of production.

 In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, *Proc. Natl. Acad. Sci.*, 81:6851-6855; Neuberger et al., 1984, *Nature*, 312:604-608; Takeda et al., 1985, *Nature*, 314:452-454) by
25 splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as
30 those having a variable region derived from a murine mAb and a human immunoglobulin constant region. Such technologies are described in U.S. Patents Nos. 6,075,181 and 5,877,397 and their respective disclosures which are herein incorporated by reference in their entirety.

35 Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, *Science* 242:423-426; Huston et al., 1988, *Proc. Natl. Acad.*

Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce single chain antibodies against NHP gene products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include, but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to a NHP can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" a given NHP, using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, 1993, FASEB J 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438). For example antibodies which bind to a NHP domain and competitively inhibit the binding of NHP to its cognate receptor can be used to generate anti-idiotypes that "mimic" the NHP and, therefore, bind and activate or neutralize a receptor. Such anti-idiotypic antibodies or Fab fragments of such anti-idiotypes can be used in therapeutic regimens involving a NHP mediated pathway.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the scope of the appended claims. All

cited publications, patents, and patent applications are herein incorporated by reference in their entirety.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising at least 24 contiguous bases of nucleotide sequence first
5 disclosed in the NHP polynucleotide described in SEQ ID NO: 1.
2. An isolated nucleic acid molecule comprising a nucleotide sequence that:
 - (a) encodes the amino acid sequence shown in SEQ ID
10 NO: 2; and
 - (b) hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NO: 1 or the complement thereof.
- 15 3. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NO: 4.

SEQUENCE LISTING

<110> LEXICON GENETICS INCORPORATED

<120> Novel Human Membrane Proteins and
Polynucleotides Encoding the Same

<130> LEX-0089-PCT

<150> US 60/165,263

<151> 1999-11-12

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1758

<212> DNA

<213> homo sapiens

<400> 1

atgcactgcc	tggggcgccga	gtacctgggtg	tttatcatgg	gctactcggt	ggtggctgga	60
gccgcgggaa	gactcctctt	tggetatgac	agctttggca	acatgtgtgg	caagaagaac	120
tcccccgtag	aagggggccc	tctttcaggg	caggacatga	ccctaaaaaa	acacgtgttc	180
tttatgaatt	cctgcaacct	ggaagtcaaa	gggtccttcc	tgtgtgttta	tagtttgaat	240
tccttcaact	acacccacag	tccaaaagca	gactcactgt	gtcccaggct	accagttcct	300
ccaagcaagt	catttccctt	atttaaccga	tgtgtccctc	aaacacctga	gtgtacttcc	360
ctatttgcac	ctgttttgat	aaatgatgtt	gacaccctcc	accgaattct	aagtgggaatc	420
atgtcgggaa	gagatacaat	ccttggcctg	tgtatcctcg	cattagcctt	gtctttggcc	480
atgatgttta	ccttcagatt	catcaccacc	cttctgggtc	acattttcat	ttcattgggt	540
attttgggat	tgttgtttgt	ctgcgggtgt	ttatgggtgg	tgtattatga	ctataccaac	600
gacctcagca	tagaattgga	cacagaaagg	gaaaatatga	agtgcgtgct	gggggttgct	660
atcgtatcca	caggcatcac	ggcagtgtcg	ctcgtcttga	tttttgttct	cagaaagaga	720
ataaaaattga	cagttgagct	tttccaaatc	acaaataaag	ccatcagcag	tgctcccttc	780
ctgctgttcc	agccactgtg	gacatttgcc	atcctcattt	tcttctgggt	cctctgggtg	840
gctgtgctgc	tgagcctggg	aactgcagga	gctgcccagg	ttatggaagg	cggccaagtg	900
gaatataagc	ccctttcggg	cattcggtac	atgtggtcgt	accatttaat	tggcctcatc	960
tggactagtg	aattcactct	tgcgtgccag	caaatgacta	tagctggggc	agtggttact	1020
tgttatttca	acagaagtaa	aaatgatcct	cctgatcatc	ccatcctttc	gtctctctcc	1080
attctcttct	tctaccatca	aggaaccrtt	gtgaaagggt	catttttaat	ctctgtgggt	1140
aggattccga	gaatcattgt	catgtacatg	caaaacgcac	tgaaagaaca	gcagcatggg	1200
gcattgtcca	ggtacctgtt	ccgatgctgc	tactgctgtt	tctggtgtct	tgacaaatac	1260
ctgctccatc	tcaaccagaa	tgcataact	acaactgcta	ttaatgggac	agatttctgt	1320
acatcagcaa	aagatgcatt	caaaatcttg	tccaagaact	caagtcactt	tacatctatt	1380
aactgctttg	gagacttcat	aatttttcta	ggaaagggtg	tagtgggtgt	tttactgtt	1440
tttggaggac	tcatggcttt	taactacaat	cgggcattcc	aggtgtgggc	agtccctctg	1500
ttattggtag	ctttttttgc	ctacttagta	gcccatagtt	ttttatctgt	gtttgaaact	1560
gtgctggatg	cacttttctc	gtgttttgct	gttgatctgg	aaacaaatga	tggatcgtca	1620
gaaaagccct	actttatgga	tcaagaattt	ctgagtttct	taaaaaggag	caacaaatta	1680
aacaatgcaa	gggcacagca	ggacaagcac	tcattaagga	atgaggaggg	aacagaactc	1740
caggccattg	tgagatag					1758

<210> 2

<211> 585

<212> PRT

<213> homo sapiens

<400> 2

Met	His	Cys	Leu	Gly	Ala	Glu	Tyr	Leu	Val	Phe	Ile	Met	Gly	Tyr	Ser
1				5						10				15	

Val	Val	Ala	Gly	Ala	Ala	Gly	Arg	Leu	Leu	Phe	Gly	Tyr	Asp	Ser	Phe
		20						25					30		
Gly	Asn	Met	Cys	Gly	Lys	Lys	Asn	Ser	Pro	Val	Glu	Gly	Ala	Pro	Leu
		35					40					45			
Ser	Gly	Gln	Asp	Met	Thr	Leu	Lys	Lys	His	Val	Phe	Phe	Met	Asn	Ser
		50				55					60				
Cys	Asn	Leu	Glu	Val	Lys	Gly	Ser	Phe	Leu	Cys	Val	Tyr	Ser	Leu	Asn
65					70					75				80	
Ser	Phe	Asn	Tyr	Thr	His	Ser	Pro	Lys	Ala	Asp	Ser	Leu	Cys	Pro	Arg
				85				90						95	
Leu	Pro	Val	Pro	Pro	Ser	Lys	Ser	Phe	Pro	Leu	Phe	Asn	Arg	Cys	Val
			100					105					110		
Pro	Gln	Thr	Pro	Glu	Cys	Tyr	Ser	Leu	Phe	Ala	Ser	Val	Leu	Ile	Asn
		115					120					125			
Asp	Val	Asp	Thr	Leu	His	Arg	Ile	Leu	Ser	Gly	Ile	Met	Ser	Gly	Arg
		130				135					140				
Asp	Thr	Ile	Leu	Gly	Leu	Cys	Ile	Leu	Ala	Leu	Ala	Leu	Ser	Leu	Ala
145					150					155					160
Met	Met	Phe	Thr	Phe	Arg	Phe	Ile	Thr	Thr	Leu	Leu	Val	His	Ile	Phe
				165				170						175	
Ile	Ser	Leu	Val	Ile	Leu	Gly	Leu	Leu	Phe	Val	Cys	Gly	Val	Leu	Trp
			180				185						190		
Trp	Leu	Tyr	Tyr	Asp	Tyr	Thr	Asn	Asp	Leu	Ser	Ile	Glu	Leu	Asp	Thr
		195					200					205			
Glu	Arg	Glu	Asn	Met	Lys	Cys	Val	Leu	Gly	Phe	Ala	Ile	Val	Ser	Thr
		210				215					220				
Gly	Ile	Thr	Ala	Val	Leu	Leu	Val	Leu	Ile	Phe	Val	Leu	Arg	Lys	Arg
225					230					235					240
Ile	Lys	Leu	Thr	Val	Glu	Leu	Phe	Gln	Ile	Thr	Asn	Lys	Ala	Ile	Ser
				245				250						255	
Ser	Ala	Pro	Phe	Leu	Leu	Phe	Gln	Pro	Leu	Trp	Thr	Phe	Ala	Ile	Leu
			260				265						270		
Ile	Phe	Phe	Trp	Val	Leu	Trp	Val	Ala	Val	Leu	Leu	Ser	Leu	Gly	Thr
		275					280					285			
Ala	Gly	Ala	Ala	Gln	Val	Met	Glu	Gly	Gly	Gln	Val	Glu	Tyr	Lys	Pro
		290				295				300					
Leu	Ser	Gly	Ile	Arg	Tyr	Met	Trp	Ser	Tyr	His	Leu	Ile	Gly	Leu	Ile
305					310					315					320
Trp	Thr	Ser	Glu	Phe	Ile	Leu	Ala	Cys	Gln	Gln	Met	Thr	Ile	Ala	Gly
				325				330						335	
Ala	Val	Val	Thr	Cys	Tyr	Phe	Asn	Arg	Ser	Lys	Asn	Asp	Pro	Pro	Asp
			340					345					350		
His	Pro	Ile	Leu	Ser	Ser	Leu	Ser	Ile	Leu	Phe	Phe	Tyr	His	Gln	Gly
		355					360					365			
Thr	Ile	Val	Lys	Gly	Ser	Phe	Leu	Ile	Ser	Val	Val	Arg	Ile	Pro	Arg
						375				380					
Ile	Ile	Val	Met	Tyr	Met	Gln	Asn	Ala	Leu	Lys	Glu	Gln	Gln	His	Gly
385					390					395					400
Ala	Leu	Ser	Arg	Tyr	Leu	Phe	Arg	Cys	Cys	Tyr	Cys	Cys	Phe	Trp	Cys
				405				410						415	
Leu	Asp	Lys	Tyr	Leu	Leu	His	Leu	Asn	Gln	Asn	Ala	Tyr	Thr	Thr	Thr
			420					425					430		
Ala	Ile	Asn	Gly	Thr	Asp	Phe	Cys	Thr	Ser	Ala	Lys	Asp	Ala	Phe	Lys
		435					440					445			
Ile	Leu	Ser	Lys	Asn	Ser	Ser	His	Phe	Thr	Ser	Ile	Asn	Cys	Phe	Gly
						455					460				
Asp	Phe	Ile	Ile	Phe	Leu	Gly	Lys	Val	Leu	Val	Val	Cys	Phe	Thr	Val
465					470					475					480
Phe	Gly	Gly	Leu	Met	Ala	Phe	Asn	Tyr	Asn	Arg	Ala	Phe	Gln	Val	Trp
				485					490					495	
Ala	Val	Pro	Leu	Leu	Leu	Val	Ala	Phe	Phe	Ala	Tyr	Leu	Val	Ala	His
			500					505				510			
Ser	Phe	Leu	Ser	Val	Phe	Glu	Thr	Val	Leu	Asp	Ala	Leu	Phe	Leu	Cys

```

      515              520              525
Phe Ala Val Asp Leu Glu Thr Asn Asp Gly Ser Ser Glu Lys Pro Tyr
      530              535              540
Phe Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg Ser Asn Lys Leu
545              550              555              560
Asn Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu Arg Asn Glu Glu
      565              570              575
Gly Thr Glu Leu Gln Ala Ile Val Arg
      580              585

```

<210> 3
 <211> 1755
 <212> DNA
 <213> homo sapiens

```

<400> 3
atgcactgcc tgggcgccga gtacctgggtg tttatcatgg gctactcggt ggtggctgga      60
gccgcgggaa gactcctctt tggctatgac agctttggca acatgtgtgg caagaagaac      120
tcccctcggtg aaggggcccc tctttcaggg caggacatga ccctaaaaaa acacgtgttc      180
ttttgaatt cctgcaacct ggaagtcaaa gggctcctcc tgtgtgttta tagtttgaat      240
tccttcaact acaccacag tccaaaagca gactcactgt gtcccaggct accagttcct      300
ccaagcaagt catttccctt atttaaccga tgtgtccctc aaacacctga gtgctactcc      360
ctatttgcatt ctgttttgat aaatgatgtt gacaccctcc accgaattct aagtgggaatc      420
atgtcgggaa gagatacaat ccttgacctg tgtatcctcg cattagcctt gtctttggcc      480
atgatgttta ccttcagatt catcaccacc ctctcgtgtc acattttcat ttcatgggtt      540
attttgggat tgttgtttgt ctgcggtgtt ttatggtggc tgtattatga ctataccaac      600
gacctcagca tagaattgga cacagaaagg gaaaatatga agtgcgtgct ggggtttgct      660
atcgtatcca caggcatcac ggcagtgtcg ctctgtctga tttttgttct cagaaagaga      720
ataaaaattga cagttgagct tttccaaatc acaataaag ccatacagcag tgctcccttc      780
ctgctgttcc agccactgtg gacatttgcc atcctcattt tcttctgggt cctctgggtg      840
gctgtgctgc tgagcctggg aactgcagga gctgcccagg ttatggaagg cggccaagtg      900
gaatataagc ccctttcggg cattcgggtac atgtggctgt accatttaat tggcctcatc      960
tggactagtg aattcatcct tgcgtgccag caaatgacta tagctggggc agtggttact      1020
tgttatttca acagaagtaa aaatgatcct cctgatcatc ccataccttc gtctctctcc      1080
attctcttct tctaccatca aggaaccrtt gtgaaagggt catttttaat ctctgtgggt      1140
aggattccga gaatcattgt catgtacatg caaaacgcac tgaaagaaca gcatggtgca      1200
ttgtccaggt acctgttccg atgctgtctac tgctgtttct ggtgtcttga caaataacctg      1260
ctccatctca accagaatgc atatactaca actgctatta atgggacaga tttctgtaca      1320
tcagcaaaag atgcattcaa aatcttgtcc aagaactcaa gtcactttac atctattaac      1380
tgcctttggag acttcataat ttttctagga aagggtgttag tgggtgtgtt cactgttttt      1440
ggaggactca tggcttttaa ctacaatcgg gcattccagg tgtgggcagt ccctctgtta      1500
ttggtagctt tttttgcta cttagtagcc catagttttt tatctgtgtt tgaaactgtg      1560
ctggatgcac ttttctgtg ttttctgtgt gatctggaaa caaatgatgg atcgtcagaa      1620
aagccctact ttatggatca agaatttctg agtttcgtaa aaaggagcaa caaattaaac      1680
aatgcaaggg cacagcagga caagcactca ttaaggaatg aggagggaac agaactccag      1740
gccattgtga gatag      1755

```

<210> 4
 <211> 584
 <212> PRT
 <213> homo sapiens

```

<400> 4
Met His Cys Leu Gly Ala Glu Tyr Leu Val Phe Ile Met Gly Tyr Ser
1              5              10              15
Val Val Ala Gly Ala Ala Gly Arg Leu Leu Phe Gly Tyr Asp Ser Phe
      20              25              30
Gly Asn Met Cys Gly Lys Lys Asn Ser Pro Val Glu Gly Ala Pro Leu
      35              40              45
Ser Gly Gln Asp Met Thr Leu Lys Lys His Val Phe Phe Met Asn Ser
      50              55              60
Cys Asn Leu Glu Val Lys Gly Ser Phe Leu Cys Val Tyr Ser Leu Asn
      65              70              75              80

```

Ser Phe Asn Tyr Thr His Ser Pro Lys Ala Asp Ser Leu Cys Pro Arg
 85 90 95
 Leu Pro Val Pro Pro Ser Lys Ser Phe Pro Leu Phe Asn Arg Cys Val
 100 105 110
 Pro Gln Thr Pro Glu Cys Tyr Ser Leu Phe Ala Ser Val Leu Ile Asn
 115 120 125
 Asp Val Asp Thr Leu His Arg Ile Leu Ser Gly Ile Met Ser Gly Arg
 130 135 140
 Asp Thr Ile Leu Gly Leu Cys Ile Leu Ala Leu Ala Leu Ser Leu Ala
 145 150 155 160
 Met Met Phe Thr Phe Arg Phe Ile Thr Thr Leu Leu Val His Ile Phe
 165 170 175
 Ile Ser Leu Val Ile Leu Gly Leu Leu Phe Val Cys Gly Val Leu Trp
 180 185 190
 Trp Leu Tyr Tyr Asp Tyr Thr Asn Asp Leu Ser Ile Glu Leu Asp Thr
 195 200 205
 Glu Arg Glu Asn Met Lys Cys Val Leu Gly Phe Ala Ile Val Ser Thr
 210 215 220
 Gly Ile Thr Ala Val Leu Leu Val Leu Ile Phe Val Leu Arg Lys Arg
 225 230 235 240
 Ile Lys Leu Thr Val Glu Leu Phe Gln Ile Thr Asn Lys Ala Ile Ser
 245 250 255
 Ser Ala Pro Phe Leu Leu Phe Gln Pro Leu Trp Thr Phe Ala Ile Leu
 260 265 270
 Ile Phe Phe Trp Val Leu Trp Val Ala Val Leu Leu Ser Leu Gly Thr
 275 280 285
 Ala Gly Ala Ala Gln Val Met Glu Gly Gly Gln Val Glu Tyr Lys Pro
 290 295 300
 Leu Ser Gly Ile Arg Tyr Met Trp Ser Tyr His Leu Ile Gly Leu Ile
 305 310 315 320
 Trp Thr Ser Glu Phe Ile Leu Ala Cys Gln Gln Met Thr Ile Ala Gly
 325 330 335
 Ala Val Val Thr Cys Tyr Phe Asn Arg Ser Lys Asn Asp Pro Pro Asp
 340 345 350
 His Pro Ile Leu Ser Ser Leu Ser Ile Leu Phe Phe Tyr His Gln Gly
 355 360 365
 Thr Ile Val Lys Gly Ser Phe Leu Ile Ser Val Val Arg Ile Pro Arg
 370 375 380
 Ile Ile Val Met Tyr Met Gln Asn Ala Leu Lys Glu Gln His Gly Ala
 385 390 395 400
 Leu Ser Arg Tyr Leu Phe Arg Cys Cys Tyr Cys Cys Phe Trp Cys Leu
 405 410 415
 Asp Lys Tyr Leu Leu His Leu Asn Gln Asn Ala Tyr Thr Thr Ala
 420 425 430
 Ile Asn Gly Thr Asp Phe Cys Thr Ser Ala Lys Asp Ala Phe Lys Ile
 435 440 445
 Leu Ser Lys Asn Ser Ser His Phe Thr Ser Ile Asn Cys Phe Gly Asp
 450 455 460
 Phe Ile Ile Phe Leu Gly Lys Val Leu Val Val Cys Phe Thr Val Phe
 465 470 475 480
 Gly Gly Leu Met Ala Phe Asn Tyr Asn Arg Ala Phe Gln Val Trp Ala
 485 490 495
 Val Pro Leu Leu Leu Val Ala Phe Phe Ala Tyr Leu Val Ala His Ser
 500 505 510
 Phe Leu Ser Val Phe Glu Thr Val Leu Asp Ala Leu Phe Leu Cys Phe
 515 520 525
 Ala Val Asp Leu Glu Thr Asn Asp Gly Ser Ser Glu Lys Pro Tyr Phe
 530 535 540
 Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg Ser Asn Lys Leu Asn
 545 550 555 560
 Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu Arg Asn Glu Glu Gly
 565 570 575
 Thr Glu Leu Gln Ala Ile Val Arg

580

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/31045

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C12N15/11 C12N15/62 C07K14/705 C07K16/28
 A61K38/17 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EMBL, STRAND, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EM_EST [Online] EMBL; ID HS784137, AC R75784, 10 June 1995 (1995-06-10) HILLIER L ET AL.: "yl21d10.r1 Soares breast 2NbHBst Homo sapiens cDNA clone IMAGE:158899 5', mRNA sequence." XP002164777 Note: 97.7% nt seq identity (99.5% ungapped) in 383 nt overlap with SEQ ID NO:1 (1-383:233-608) the whole document</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1-3

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

20 April 2001

Date of mailing of the international search report

26.04.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/31045

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EM_EST [Online] EMBL; ID HSZZ34483, AC AA329432, 18 April 1997 (1997-04-18) ADAMS M D ET AL.: "EST33039 Embryo, 12 week II Homo sapiens cDNA 5' end." XP002164778 Note: 98.1% nt seq identity (99.4% ungapped) in 313 nt overlap with SEQ ID NO:1 (1-310:1030-1342) the whole document</p> <p>---</p>	1-3
X	<p>DATABASE EM_EST [Online] EMBL; ID AA809640, AC AA809640, 16 February 1998 (1998-02-16) STRAUSBERG R, NATIONAL CANCER INSTITUTE, CANCER GENOME ANATOMY PROJECT (CGAP): "nk96a12.s1 NCI CGAP Co3 Homo sapiens cDNA clone IMAGE:1028638 3' similar to WP:F35C8.7 CE04499 ;, mRNA sequence." XP002164779 Note: 99.7% nt seq identity (99.7% ungapped) in 310 nt overlap with SEQ ID NO:1 (133-442:1758-1449) the whole document</p> <p>---</p>	1-3
A	<p>WO 98 27205 A (GENETICS INST) 25 June 1998 (1998-06-25) Note: 53.1% aa seq identity (53.3% ungapped) in 277 aa overlap of SEQ ID NO:21 with SEQ ID NO:2 (3-278:283-559) page 15, line 1 -page 16, line 22 page 26, line 15 -page 27, line 3 page 83-85 claims 35,36</p> <p>---</p>	1-3
A	<p>DATABASE SWALL [Online] SWISSPROT; ID Q9Y332, AC Q9Y332, 1 November 1999 (1999-11-01) ROWEN L ET AL.: "NG22 (Sequence of the human major histocompatibility complex class III region)" XP002164780 Note: 28.9% aa seq identity (34.6% ungapped) in 547 aa overlap with SEQ ID NO:2 (166-702:93-559) the whole document</p> <p>---</p> <p style="text-align: center;">-/--</p>	1-3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/31045

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	AGUADO B ET AL: "Genes of the MHC class III region and the functions of the proteins they encode" HLA MHC: GENES, MOL. FUNCT. EDITOR(S): BROWNING, MICHAEL; MCMICHAEL, ANDREW. PUBLISHER: BIOS SCIENTIFIC PUBLISHERS, OXFORD, UK, 1996, pages 39-75, XP000998273 page 39, line 1 -page 40, line 6 ---	
A	THE MHC SEQUENCING CONSORTIUM: "Complete sequence and gene map of a human major histocompatibility complex" NATURE, vol. 401, no. 6756, 28 October 1999 (1999-10-28), pages 921-923, XP002164864 abstract ---	
P,X	WO 99 63088 A (BAKER KEVIN ;CHEN JIAN (US); GENENTECH INC (US); YUAN JEAN (US); G) 9 December 1999 (1999-12-09) Note: 99.8% aa seq identity (99.8% ungapped) in 445 aa overlap of PR01115 with SEQ ID NO:2 (1-445:141-585) page 1, line 6 -page 2, line 6 page 14, paragraph 43 page 108-109, paragraph 43 page 324, paragraph 43 example 46 figures 110,111 -----	1-3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/31045

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9827205 A	25-06-1998	AU 5706198 A EP 0958360 A	15-07-1998 24-11-1999
WO 9963088 A	09-12-1999	AU 4328699 A AU 2212299 A WO 9935170 A	20-12-1999 26-07-1999 15-07-1999